MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Research Report 506

EXPECTED UTILITY, PENALTY FUNCTIONS,
AND DUALITY IN STOCHASTIC
NONLINEAR PROGRAMMING

by

A. Ben-Tal*
M. Teboulle*

CENTER FOR CYBERNETIC STUDIES

The University of Texas
Austin, Texas 78712

DTIC
ELECTE
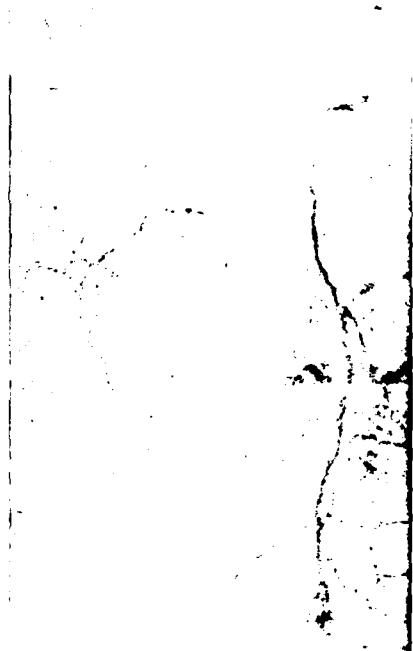S   D
MAY 2 1 1985

B

Research Report 506

EXPECTED UTILITY, PENALTY FUNCTIONS, AND DUALITY IN STOCHASTIC NONLINEAR PROGRAMMING
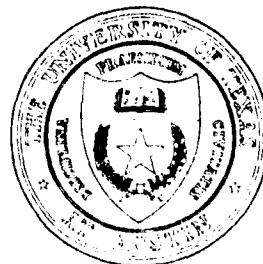
by

A. Ben-Tal*
M. Teboulle*

June 1984
Revised February 1985

*The Technion, Israel Institute of Technology, Haifa, Israel

DTIC
ELECTE
MAY 2 1 1985
S   D

B

CENTER FOR CYBERNETIC STUDIES

A. Charnes, Director
College of Business Administration 5.202
The University of Texas at Austin
Austin, Texas 78712-1177
(512) 471-1821

ABSTRACT

We consider nonlinear programming problems with stochastic constraints.
The Lagrangian corresponding to such problems has a stochastic part, which
in this work is replaced by its certainty equivalent (in the sense of expected
utility theory). It is shown that the deterministic surrogate problem thus
obtained, contains a penalty function which penalized violation of the con-
straints in the mean. The dual problem is studied (for problems with sto-
chastic righthand sides in the constraints) and a comprehensive duality theory
is developed by introducing a new certainty equivalent concept, which possesses,
for arbitrary utility functions, some of the properties that the classical cer-
tainty equivalent retains only for the exponential utility.

KEY WORDS

Stochastic Programming
Duality in Nonlinear Programming
Minmax Theorems
Convex Functions
Expected Utility
Risk Aversion
Certainty Equivalent
Penalty Functions

1. INTRODUCTION

Consider the non-linear programming problem

(P) $\inf g_0(x)$

    s.t $g_i(x,b) \leqslant 0$ $i \in I = \{1,\ldots,m\}$

where $x \in \mathbb{R}^n$ is the decision vector; $b \in \mathbb{R}^k$ is a given fixed vector of parameters, $g_0: \mathbb{R}^n \to \mathbb{R}$ and $g_i: \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}$ are given real-valued functions. We will sometimes use also the vector notation $g(x,b) = (g_1(x,b),\ldots,g_m(x,b))^{\mathsf{T}}$. Let the feasible set be denoted by:

$$S = \{x: g_i(x,b) \leqslant 0 \quad \forall i \in I\} \ .$$

A fundamental approach to solve (P) is to replace it by an unconstrained problem of the form:

$$\inf\{g_0(x) + P(x)\}$$

where $P(x)$ is a penalty function prescribing a "high cost" for violation of the constraints (see e.g., Fiacco and McCormick [1968], Luenberger [1973]).

An ideal penalty function for problem (P) is:

$$P(x) = \begin{cases} 0 & \text{if } x \in S \\ +\infty & \text{otherwise} \end{cases}$$

In fact such a penalty is implicitly embedded in the Lagrangian associated with problem (P):

$$L_b(x,y) = g_0(x) + y^{\mathsf{T}} g(x,b) \ .$$

Indeed, (P) can be reformulated as a <u>saddle-function</u> problem:

$$\inf_{x \in \mathbb{R}^n} \sup_{y \geq 0} L_b(x,y) \tag{1}$$

Which is clearly equivalent to the <u>unconstrained</u> optimization problem:

$$\inf_x \{g_o(x) + P(x)\} \tag{2}$$

where

$$P(x) = \sup_{y \geq 0} y^T g(x,b) . \tag{3}$$

Since the supremum is equal to zero if $x \in S$, and to infinity if $x \notin S$, $P(x)$, as defined in (3), is an ideal penalty for problem (P).

Assume now, and henceforth in this paper, that the parameter b is a <u>random vector</u>, with known distribution function F_b, and support $B \subset \mathbf{R}^k$. Problem (P) will be now referred to as <u>Stochastic Program</u> (SP). The penalty function, just defined in (3), is now the solution of a (linear) programming problem with a <u>stochastic objective function</u>. For such problems the classical economic theory approach of decision under uncertainty, is to replace the stochastic objective function by its expected utility $Eu(y^T g(x,b))$, where u is a Von-Neumann-Morgenstern utility function, and E denotes the expectation operator with respect to b. Adopting this approach here, we replace the stochastic objective function $y^T g(x,b)$ by its <u>certainty equivalent</u> $u^{-1} Eu(y^T g(x,b))$; thus replacing the penalty function (3) by:

$$P_u(x) = \sup_{y \geq 0} u^{-1} Eu(y^T g(x,b)) \tag{4}$$

where u^{-1} is the inverse of u.

The use of $u^{-1}Eu(\cdot)$ in (4), rather than $Eu(\cdot)$, is appropriate since in (2) we must add terms with comparable units (not \$ + utiles). Using the u-penalty (4) in problem (2), the stochastic program (SP) is thus replaced by what we will call the Certainty Equivalent Primal problem:

$$(CE-P) \qquad \inf_{x}\{g_o(x) + P_u(x)\}$$

In the first part of this paper we study problem (CE-P) and show its relevance as a penalty approach to treat programming problems with nonlinear stochastic constraints. A related approach for the case of Stochastic Linear programs is given in Ben-Tal and Teboulle [1984]. The study of (CE-P) relies on the properties of the u-penalty $P_u(x)$, and these will be derived in §4 using preliminary results from the two preceding sections: Section 2, which summarizes some basic facts from utility theory, and Section 3, in which we prove a convexity result on the certainty equivalent functional:

$$v(y) = u^{-1}Eu(y^TZ) \ .$$

The latter result is needed to demonstrate that P_u, given in (4), is the optimal value of a concave program.

In the deterministic case, the representation of (P) as the saddle function problem (1), is also the source of obtaining a dual problem associated with (P), namely:

$$(D) \qquad \sup_{y \geqslant 0} \inf_{x} L_b(x,y) \ .$$

In the stochastic case, it is then natural to study the nature of the dual problem corresponding to (CE-P). This will be carried out

in the second part of the paper (§ 5-8) for the important special case of problems with <u>stochastic righthand side</u>:

(SP-RHS) $\inf\{g_o(x): g(x) \geqslant b\}$.

This question was addressed recently, for exponential utilities, in Ben-Tal [1984]. It was shown there that the dual problem of (CE-P) consists of maximizing the certainty equivalent of the Lagrangian dual function:

$$h_b(y) = \inf_x L_b(x,y) \ ,$$

i.e., the dual problem is

(CE-D) $\max_{y \geqslant 0} u^{-1} Eu(h_b(y))$.

This result is recovered here in Section 5. However, for arbitrary utilities, such a duality result does not hold; this is due mainly to the non-additivity of the certainty equivalent for non-exponential utilities. Therefore, we suggest, in Section 6 of this paper, a new type of a certainty equivalent functional, which possesses, for <u>arbitrary utilities</u>, many of the properties that the classical certainty equivalent possesses <u>only</u> for exponential utilities[+]. The appropriateness of the new certainty equivalent in defining a corresponding u-penalty function, and its use in treating stochastic programs is discussed in Section 7. A complete duality theory for (CE-P) is then obtained in Section 8.

[+] For the special role of exponential utility in economic analysis see e.g., Bamberg and Spremann [1981].

2. SOME PRELIMINARIES ON UTILITY FUNCTIONS

In their classical work, von-Neumann and Morgenstern [1947] have developed a set of axioms concerning preferences over probability distributions. Under these axioms, a random variable X is preferred against the random variable Y if and only if there exists a real valued function u, called a <u>utility function</u>, unique up to a monotone increasing affine transformation, such that $Eu(X) \geqslant Eu(Y)$, where E denotes the mathematical expectation. A decision maker is called <u>risk averter</u> if E(X) is preferred against X for all random variables X, i.e.,

$$u[E(X)] \geqslant Eu[X]$$

The latter is the Jensen inequality for $u(\cdot)$, and therefore equivalent to the concavity of the utility function u. Throughout this paper we shall deal with the class U of strictly increasing concave utilities. Further, denote by U_k the class of strictly increasing, strictly concave; k-times continuously differentiable functions, more precisely:

$$U_1 = \{u \in C^1 : u' > 0 \text{ and } u \text{ strictly concave}\}$$
$$U_k = \{u \in C^{(k)} : u' > 0 \text{ and } u'' < 0\} \qquad k \geqslant 2 .$$

Following Pratt [1964] and Arrow [1971], the <u>measure of local risk aversion</u> at the point $t \in \mathbb{R}$ for $u \in U_2$, is defined by:

$$r(t) = - \frac{u''(t)}{u'(t)} . \tag{5}$$

Three basic properties of r(t) for $u \in U_2$ are summarized below:

a) r(t) is well-defined and $r(t) > 0 \quad \forall t \in \mathbb{R}$.

b) r(t) is <u>invariant</u> with respect to any positive affine transformation of the utility function u(t) .

c) Given $r(t)$, the utility function $u(t)$ is uniquely determined (up to a positive affine transformation) by:

$$u(t) = \int^t \exp(-\int^s r(\sigma)d\sigma)ds$$

□

For a given $u \in U$, the inverse function u^{-1} exists and is a strictly increasing convex function. The Certainty Equivalent of a random variable X is defined by

$$C(X) = u^{-1}Eu(X) \quad .$$

It is the sure amount for which the decision maker remains indifferent to a gamble yielding X, i.e.,

$$u(C(X)) = Eu(X) \quad .$$

The following properties of the certainty equivalent are immediate consequences of its definition.

Proposition 1   Let X be a random variable and $u \in U$ then:

(a) $C(w) = w \qquad \forall w \in \mathbb{R}$.

(b) $C(X)$ is invariant to affine transformation in u.

(c) $C(X) \leq E[X]$ with equality for all X if and only if u is linear.

□

Looking back to the definition of the u-penalty function:

$$P_u(x) = \sup_{y \geq 0} \{C(y^T g(x,b)) = u^{-1}Eu[y^T g(x,b)]\} \tag{4}$$

we see from Proposition 1 that it possesses two desirable features:

(a)   In the case where b is a deterministic vector of parameters, the original problem (P) is recovered from (CE-P) ;

(b)   The penalty function is invariant to affine transformations of u.

There is, however, some difficulty associated with the objective function in (4):

$$Q(x;y) \triangleq C[y^T g(x,b)] = u^{-1} E u (y^T g(x,b))$$

in term of which P_u is computed. As a function of y, $Q(x;y)$ is a _convex increasing_ transformation (u^{-1}) of the _concave_ function $E u (y^T g(x,b))$, thus in general it is not guaranteed that $Q(x;\cdot)$ is concave.

In the next section we characterized the utilities for which $y \rightarrow Q(x;y)$ is a concave function, for any random vector b. The characterization is given in term of the Arrow-Pratt risk aversion measure r.

3. CONCAVITY OF THE CERTAINTY EQUIVALENCE FUNCTIONAL

Let Z be a random vector in $\mathbb{R}^m$, and for $u \in U$ define the _certainty equivalence functional_ by

$$v(y) = u^{-1} E u (y^T Z) .$$

We further assume that $E u (y^T Z) < +\infty$ $\forall y \in \mathbb{R}^m$.

Let $\phi = u^{-1}$, we next define a function of two variables h, which plays a central role in the proof of the main result of this section:

$$h(x_1, x_2) := \frac{\phi(x_1) - \phi(x_2)}{\phi'(x_2)} .$$

Note that since $\phi' > 0$, $h(x_1, x_2)$ is well defined for all x_1, x_2 in the range of u and is a twice continuously differentiable function for $u \in U_3$. The convexity of $h(x_1, x_2)$ will be now characterized

by showing that its Hessian matrix $\nabla^2 h$ is positive semi-definite, i.e.,

$$\forall\ d \in \mathbb{R}^2 \qquad d^T \nabla^2 h d \geqslant 0.$$

<u>Lemma 1</u> - Let $u \in U_3$, then the function $h(x_1, x_2)$ is convex if and only if $\frac{1}{r(t)}$ is concave.

<u>Proof</u>: The Hessian matrix of h will be positive semi-definite if and only if:

$$D(x,y) := x^2 h_{11} + y^2 h_{22} + 2xy h_{12} \geqslant 0 \quad \text{for any } x,y \text{ not both zero}$$

where $h_{ij} = \frac{\partial^2 h}{\partial x_i \partial x_j}$ $\quad i,j = 1,2$.

Computing the partial derivatives of $h(x_1, x_2)$ we obtain:

$$h_{11} = \frac{\phi''(x_1)}{\phi'(x_2)} \quad ; \quad h_{12} = -\frac{\phi'(x_1)\phi''(x_2)}{\phi'^2(x_2)}$$

$$h_{22} = \frac{\phi''(x_2)}{\phi'(x_2)} + h(x_1,x_2)\left\{ \frac{2\phi''^2(x_2) - \phi'''(x_2)\phi'(x_2)}{\phi'(x_2)^2} \right\}$$

Now $h_{11} > 0$ since $u \in U_2$, hence $D(x,y)$ can be rewritten as

$$D(x,y) = h_{11}\left(x + \frac{h_{12}y}{h_{11}}\right)^2 + \left(h_{22} - \frac{h_{12}^2}{h_{11}}\right) y^2 ,$$

and is non-negative for all (x,y) if and only if :

$$\Delta_h := h_{22} - \frac{h_{12}^2}{h_{11}} \geqslant 0 \quad . \tag{6}$$

In terms of $\phi(t)$, the risk aversion function defined in (5) is:

$$r(t) = -\frac{[\phi^{-1}(t)]''}{[\phi^{-1}(t)]'} \quad .$$

By the inverse function theorem we have:

$$u'[\phi(t)] = \frac{1}{\phi'(t)} \quad ; \quad u''[\phi(t)] = -\frac{\phi''(t)}{\phi'^3(t)}$$

$$u'''[\phi(t)] = \frac{3\phi''^2(t) - \phi'''(t)\phi'(t)}{(\phi'(t))^5}$$

and hence:

$$r[\phi(t)] = \frac{\phi''(t)}{\phi'^2(t)} \quad ; \quad r'[\phi(t)] = \frac{\phi'(t)\phi'''(t) - 2\phi''^2(t)}{\phi'^4(t)} \quad .$$

Now, put $t_1 \stackrel{\Delta}{=} \phi(x_1)$ and $t_2 \stackrel{\Delta}{=} \phi(x_2)$; computing Δ_h, and using the fact that $\phi' > 0$, we obtain after some algebraic manipulations that (6) is equivalent to:

$$r^2(t_2) \leqslant r(t_1)r(t_2) - (t_1-t_2)r(t_1)r'(t_2) . \tag{7}$$

Dividing (7) by $r(t_1)r^2(t_2)$, which is strictly positive, we have :

$$\frac{1}{r(t_1)} \leqslant \frac{1}{r(t_2)} - (t_1-t_2)\frac{r'(t_2)}{r^2(t_2)}$$

and this is exactly the gradient inequality for $1/r(t)$, which characterizes its concavity.

$\square$

Theorem 1: Let $u \in U_3$, then the function:

$$v(y) = u^{-1}E u (y^T Z)$$

is concave for any random vector Z, if and only if $\frac{1}{r(t)}$ is concave.

Proof:
$v(y)$ is concave if and only if it satisfies the gradient inequality, i.e.,using again the notation $\phi = u^{-1}$, and observing that $\phi' > 0$:

Clearly then $\psi'(p) \leqslant 0$, i.e., ψ is decreasing; moreover $u' > 0$ and $\alpha(p)$ is strictly increasing (the composition of two strictly decreasing functions) and therefore $\alpha'(p) > 0$, and thus by (21) $\psi'' > 0$ showing that ψ is strictly convex.

Finally, since $u'(0) = 1$, $\alpha(1) = 0$ and thus:

$$\psi(1) = u(0) = 0.$$

□

Let Z be a random vector in $\mathbb{R}^m$ and for $u \in U_N$ we define now the <u>New Certainty Equivalent</u> (NCE) <u>functional</u> by $w: \mathbb{R}^m \to \mathbb{R}$

$$w(y) = \sup_{\eta \in \mathbb{R}} \{\eta + Eu(y^T Z - \eta)\} . \tag{22}$$

The next result shows that $w(y)$ is a concave function for <u>any</u> utility function $u \in U_N$ (compare with Theorem 1, § 3).

<u>Proposition 5</u> - The function $w: \mathbb{R}^m \to \mathbb{R}$ defined in (22) is <u>concave</u> for any $u \in U_N$ and for any random vector Z.

<u>Proof</u>: The function w can be rewritten as:

$$w(y) = - \inf_{\eta \in \mathbb{R}} F(\eta,y) ,$$

where $F: \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$ is defined by : $F(\eta,y) = - \eta - Eu(y^t Z - \eta)$ since u is concave, it is easily shown that $F(\eta,y)$ is (jointly) convex, hence $w(y)$ is concave (see Rockafeller [1974], Theorem 1).

□

We shall derive now an explicit form for $w(y)$.

<u>Proposition 6</u> - For any $u \in U_N$,

$$w(y) = \eta(y) + Eu(y^T Z - \eta(y)) \tag{23}$$

we get:

$$M - E u (X - \eta) \geqslant (M - u(L - \eta)) \Pr\{M - u(X - \eta) \geqslant M - u(L - \eta)\} . \qquad (20)$$

Since u is strictly increasing:

$$\Pr\{M - u(X - \eta) \geqslant M - u(L - \eta)\} = P\{X \leqslant L\} = p , \text{ and (20) becomes}$$

$$\eta + E u (X - \eta) \leqslant \eta + p u (L - \eta) + M(1 - p)$$

thus

$$S(X) \leqslant M(1 - p) + \sup_{\eta \in \mathbb{R}} \{\eta + p u (L - \eta)\} .$$

Differentiating the supremand with respect to η and equating to zero:

$$u'(L - \eta) = \frac{1}{p}$$

This equation indeed has a solution η for all $0 < p \leqslant 1$; to verify this
note that $\{u'(0) = 1, u' \text{ is decreasing, and assumption A2}\}$ imply

$$\text{Range } u' \supset [1, +\infty) \supset \{\frac{1}{p}\}, \quad \forall \ 0 < p \leqslant 1.$$

in fact, the optimal η is

$$\eta = L - (u')^{-1}(\frac{1}{p}),$$

where $(u')^{-1}$ denotes the inverse of u'. Inequality (19) thus follows.

Now $\psi(p)$ can be written as:

$$\psi(p) = M - p(M - u(\alpha(p))) - \alpha(p)$$

with $\alpha(p) = (u')^{-1}(\frac{1}{p})$, and the expression for its derivative is

$$\psi'(p) = -(M - u(\alpha(p))) + \alpha'(p)(p u'(\alpha(p)) - 1).$$

Note that in the latter the second term is zero since $u'(\alpha(p)) = \frac{1}{p}$, so

$$\psi'(p) = -(M - u(\alpha(p)))$$

$$\psi''(p) = \alpha'(p)u'(\alpha(p)) \qquad (21)$$

The next result gives an upper bound for S(X), which will be useful in proving later theorems 6,7, and 8. These results need additional assumptions on the utility function $u:(a,b) \to \mathbb{R}$ $(-\infty \leqslant a < b \leqslant +\infty)$

A1. u is bounded above:

$$\exists \quad M: \quad u(t) \leqslant M \; \forall \; t \in (a,b)$$

A2.[(*)] $\lim_{t \to a} u'(t) = +\infty$

Note that A2 holds if u is <u>essentially</u> <u>smooth</u> in (a,b) (see e.g. Rockafeller [1970] pg. 251); in particular it holds for all the classical utilities (log, power and exponential utilities). We denote the class of utilities $u \in U_N$ satisfying A1 and A2 by U_N^A.

<u>Lemma 2</u>:   Let $u \in U_N^A$ and let $L > x_{min}$. Denote

$$p = Pr(X \leqslant L)$$

then

$$S(X) \leqslant \psi(p) + L \tag{19}$$

where

$$\psi(p) = M(1-p) + pu((u')^{-1}(\tfrac{1}{p})) - (u')^{-1}(\tfrac{1}{p}) \; .$$

Moreover ψ is a strictly convex decreasing function in (0,1], with $\psi(1) = 0$

<u>Proof</u>:   Let $\gamma(t) = M - u(t-\eta)$ and $k = M - u(L-\eta)$. Then, since $u \in U_N^B$ $\gamma(\cdot) \geqslant 0$ and $k \geqslant 0$. Applying the general Tchebychef Inequality (see e.g., Mood et al. [1974]):

$$E(\gamma(X)) \geqslant k \; Pr(\gamma(X) \geqslant k)$$

(*) We acknowledge here the contribution of the anonymous referee in pointing out the necessity of this assumption.

Proof: (a) By definition, $S(w) = \sup_{\eta \in \mathbb{R}} \{\eta + u(w-\eta)\}$, equating the derivative of the supremand to zero we obtain $u'(w-\eta) = 1$, hence since $u'(0) = 1$ and u' is strictly decreasing the supremum is attained at $\eta = w$ and its value is then $S(w) = w$.

(b) By (17) we have:

$$S(X) = \sup_{\eta \in \mathbb{R}} \{\eta + Eu(X-\eta)\} < \sup_{\eta \in \mathbb{R}} \{\eta + E(X-\eta)\} = E(X)$$

(c) For any $\eta \in R$ we have $X-\eta \geq x_{min}-\eta$, then the result follows from (a).

(d) By definition

$$S(X+w) = \sup_{\eta \in \mathbb{R}} \{\eta + Eu(X+w-\eta) \ ,$$

hence with $\hat{\eta} = \eta-w$, one obtains:

$$S(X+w) = \sup_{\hat{\eta} \in \mathbb{R}} \{\hat{\eta}+w+Eu(X-\hat{\eta})\} = w + S(X).$$

(e) Let $u(t) = 1-e^{-t}$ be a normalized utility function, then

$$S(X) = \sup_{\eta \in \mathbb{R}} \{\eta+1-Ee^{-(X-\eta)}\}, \quad \text{which by simple calculus}$$

gives $S(X) = -\log Ee^{-X}$. On the other hand, since $u \in U_N$, u^{-1} exists and a little algebra shows that $C(X) = u^{-1}Eu(X) = -\log Ee^{-X}$; hence for exponential utility $C(X) = S(X)$.

$\square$

Example: Let X be a random variable with $X_{max} \leq 1$. Then for the (normalized) quadratic utility $u(t) = t - t^2/t$, $(t \leq 1)$ a direct computation of (18) yields:

$$S(X) = \mu - \frac{1}{2} \sigma^2$$

where μ is the mean of X, and σ^2 its variance. For this case the classical certainty equivalent is

$$C(X) = 1 - \sqrt{(\mu-1)^2 + \sigma^2} \quad .$$

$\square$

Therefore, the new certainty equivalent can be viewed as an integration of the expected utility principle and the two stage approach For the latter, see Dantzig [1955], Dantzig and Madansky [1961], Walkup and Wets [1967], Mangasarian and Rosen [1964] and the recent surveys of Dempster [1980], Kall [1982] and Wets [1983].

In the sequel we frequently interchange integration and differentiation, e.g.: $\frac{d}{dt} Eu(t+X) = E \frac{d}{dt} u(t+X)$. For this to hold it suffices to assume that $u'(.)$ is continuous and $Eu'(.) < \infty$. See Bourbaki [1958], pg. 99.

Let the support of the random variable X be $[x_{min}, x_{max}]$. The supremum in (18) is attained at the point η^* satisfying $Eu'(X-\eta^*) = 1$, and since u' is decreasing then: $1 \geq u'(x_{max}-\eta^*)$ and $1 \leq u'(x_{min}-\eta^*)$, hence together with $u'(0) = 1$ this shows that $\eta^* \in [x_{min}, x_{max}]$. In particular, for random variables with compact support, the supremum in (18) is attained. The appropriateness of $S(X)$ as a certainty equivalent measure is further supported by its basic properties which are collected below.

<u>Theorem 4</u> - For any utility function $u \in U_N$, a random variable X and a constant w:

(a) [Constancy] $S(w) = w$.

(b) [Risk aversion] $S(X) < E(X)$, $\forall$ nondegenerate random variables X.

(c) [Lower bound] If X is bounded below by x_{min} then

$S(X) \geq x_{min}$.

(d) [Additivity] $S(X+w) = S(X) + w$

(e) [Exponential-case] For the (Normalized) exponential utility function, the new certainty equivalent coincides with the classical, i.e.,

$S(X) = C(X)$.

Throughout the rest of this paper we will consider the class U_N of normalized utilities:

$$u \in U_N = \{u \in U_2: u(0) = 0, u'(0) = 1\} \quad .$$

Note that for every $u \in U_N$:

$$u(x) > 0 \quad \text{for} \quad x > 0 \quad \text{since} \quad u \quad \text{is increasing} \tag{16}$$

$$u(x) < x \quad \text{for all} \quad x \neq 0 \quad \text{since} \quad u \quad \text{is strictly concave,} \tag{17}$$

thus a normalized utility can be interpreted as a discount function, and hence, the "present value" of a future (uncertain) income Y is $Eu(Y)$.

Suppose that the decision maker, expecting a future (uncertain) income of X dollars, can consume part of X at present. If he chooses to consume η dollars, the resulting present value of X is then $\eta + Eu(X-\eta)$. Thus the sure (present) value of X, denoted $S(X)$ is the result of an optimal allocation of X between present and future consumption, i.e.,

$$S(X) = \sup_{\eta \in \mathbb{R}} \{\eta + Eu(X-\eta)\} \tag{18}$$

which is our new certainty equivalent. In fact, $S(X)$ can be written alternatively as:

$$S(X) = \sup_{\eta \in \mathbb{R}} \{\eta + E \sup_y \{u(y): \eta + y \leq X\}\}$$

and thus it is the value resulting from applying a two-stage approach ("here and now") to the stochastic program:

$$\sup \{\eta: \eta \leq X\} \quad .$$

Ben-Tal [1984]. In particular, for the important special case of stochastic rhs programs, we recover here the result showing that the dual problem of (CE-P) is equivalent to Expected utility maximization of the classical Lagrangian dual function of (SP). In the rest of the paper we aim at generalizing this duality relation for arbitrary utilities. For this purpose we introduce in the next section, a new type of certainty equivalence.

6. THE NEW CERTAINTY EQUIVALENT

In this section we introduce a new-certainty equivalent in terms of which a new u-penalty $\Pi_u(x)$ is constructed, maintaining similar properties to $P_u(x)$, as well as producing duality results for (RHS) programs for general utilities. The penalty properties of $P_u(x)$ rely essentially on the following basic properties of $C(\cdot)$: (see Proposition 1, Section 2)

(a) $C(w) = w$ $\forall$ constant w

(b) $C(X) \leqslant E(X)$ for any random variable X and $u \in U$.

On the other hand general duality results for the (RHS) case rely heavily on the additivity property of $C(\cdot)$, (Proposition 4)

(c) $C(X + w) = C(X) + w$ $\forall$ fixed w

which is valid, for arbitrary random variable X, only for exponential utilities.

The three properties (a)-(c) will serve as guidelines to define our new certainty equivalent.

deterministic primal (CE-P) is:

$$\inf_x \{g_0(x) + p \sum_{k=1}^m P_u^k(x)\} \quad .$$

Proof: The result follows immediately from the fact that in the case of exponential utility, the certainty equivalent in terms of which P_u is defined, is additive (see Bamberg and Spremann [1981], Theorem 4). □

The certainty equivalent C_α defined in (12) has the following property of additivity:

Proposition 4: Let u be an exponential utility function, then for any constant $w \in \mathbb{R}$ and any random variable X

$$C_\alpha(X + w) = C_\alpha(X) + w .$$ □

Consider now the saddle function:

$$K(x,y) = g_0(x) - \frac{1}{\alpha} \log E e^{-\alpha y^T g(x,b)}$$

Using the u-penalty, given in (13), we see that the deterministic primal (CE-P) becomes for exponential utility

$$\text{(CE-P)} \quad \inf_x \sup_{y \geq 0} K(x,y) \tag{14}$$

Recalling the definition of the Lagrangian $L_b(x,y)$ for the original problem (SP), and using Proposition 4, (14) is easily shown to be:

$$\text{(CE-P)} \quad \inf_x \sup_{y \geq 0} u^{-1} E u L_b(x,y) \tag{15}$$

The u-penalty (13) for an exponential utility, and the resulting min-max representation (15) of (CE-P), were studied recently by

gives the following <u>nonsmooth</u> optimization problem:

$$\text{(AP)} \qquad \inf \{ g_0(x) + \frac{1}{2r_0} \sum_{i=1}^{m} \frac{1}{\sigma_i^2(x)} [\max(0, m_i(x))]^2 \}$$

5. THE CASE OF EXPONENTIAL UTILITY

In this section we confine attention to the important special class of constant risk aversion utility functions, i.e., utilities with $r(t) = \frac{1}{b} = \alpha > 0$. This corresponds exactly (up to a positive affine transformation) to exponential functions, $u(t) = -e^{-\alpha t}$.

The associated certainty equivalent, for a random X is then:

$$C_\alpha(X) = -\frac{1}{\alpha} \log E e^{-\alpha X} \qquad \forall \alpha > 0 \qquad (12)$$

and the corresponding u-penalty function is given by:

$$P_u(x) = \sup_{y \geqslant 0} C_\alpha(y^T g(x,b)) \qquad (13)$$

The next result shows that for (SP) with independent constraints (see Section 4, Definition 1), the joint constraints penalty $P_u(x)$ defined in (4) is <u>additive</u>, i.e., the sum of the penalties $\{P_u^k(x): k \in I\}$ for individual constraints.

<u>Theorem 3</u>: Let u be an exponential utility function

$$u(t) = a - be^{-t/p} \qquad (p > 0, \ b > 0, \ a \in \mathbb{R})$$

For independent constraints, P_u is given by:

$$P_u(x) = \sum_{k=1}^{m} P_u^k(x)$$

where $P_u^k(x) = p \sup_{y_k \geqslant 0} \{-\log E e^{-y_k g_k(x,b_k)}\}$ and then the corresponding

Using the approximation in (CE-P), one obtains an approximate Problem (AP):

$$\text{(AP)} \qquad \inf_x \{g_o(x) + \hat{P}_u(x)\} \quad .$$

This representation can be further simplifed for <u>uncorrelated constraints</u> and in particular for <u>independent constraints</u>.

<u>Definition 1</u>:   We say that $\{g_i(x,b) \leqslant 0 \quad i \in I\}$ are <u>uncorrelated [independent] constraints</u> if the components $\{b_i\}_{i \in I}$ are uncorrelated (independent] random variables and if for each i, the i-th constraint depends only on b_i, i.e., $k = m$ and $g_i(x,b) = g_i(x,b_i)$ $\forall i \in I$. In this case the variance covariance matrix is:

$$\Sigma_{ij}(x) = \begin{cases} \sigma_i^2(x) & i = j \\ \\ 0 & i \neq j \end{cases}$$

where $\sigma_i^2(x)$ is the variance of $g_i(x,b_i)$.

The second order approximation of the u-penalty reduces to

$$\hat{P}_u(x) = \sup_{y \geqslant 0} \sum_{i=1}^{m} [y_i m_i(x) - \frac{1}{2} r_o y_i^2 \sigma_i^2(x)]$$

where $m_i(x) = E_{b_i} g_i(x,b_i)$. The latter is a maximization of a separable function, which can be carried out analytically, to obtain:

<u>Proposition 3</u>:   For (SP) with uncorrelated constraints, a second order approximation of $P_u(x)$ is

$$\hat{P}_u(x) = \frac{1}{2r_o} \sum_{i=1}^{m} \frac{1}{\sigma_i^2(x)} [\max(0, m_i(x))]^2$$

□

Thus, for uncorrelated constraints, an explicit representation of (AP)

If we define the u-penalty for the k-th constraint in the natural way as:

$$P_u^k(x) = \sup_{y_k \geq 0} u^{-1} E u (y_k g_k(x,b)) \qquad (11)$$

we get the following <u>monotonicity</u> property:

<u>Proposition 2:</u> If x^1 is less feasible than x^2, for the k-th constraint, then

$$P_u^k(x^1) \geq P_u^k(x^2) .$$

$\square$

We derive now quadratic approximation for the u-penalty function $P_u(x)$, for $u \in U_2$. First denote:

$$m(x) = Eg(x,b)$$

$$\Sigma(x) = cov(g(x,b)) .$$

It can be shown by direct differentiation of the certainty equivalent functional

$$v(y) = u^{-1} E u (y^T g(x,b))$$

that

$$\nabla v(0) = m(x)$$

$$\nabla^2 v(0) = -r_o \cdot \Sigma(x) \quad \text{with} \quad r_o := -\frac{u''(0)}{u'(0)} > 0 .$$

Moreover, we know that $v(0) = 0$, hence a second order Taylor-expression of $v(y)$ in (4) yields the following approximation of $P_u(x)$:

$$\hat{P}_u(x) = \sup_{y \geq 0} \{y^T m(x) - \frac{1}{2} r_o y^T \Sigma(x)y\}.$$

Note that, since $\Sigma(x)$ is positive semi-definite, the approximate u-penalty $\hat{P}_u(x)$ is the optimal value of a <u>concave quadratic program</u> with only nonnegativity constraints.

Then we have $Q_k(x,0) = 0$ and using the chain rule, with u' being continuous:

$$\frac{\partial}{\partial y_k} Q(x,y_k)\Big]_{y_k=0} = Eg_k(x,b) > 0 .$$

Hence, there exists $\hat{y}_k > 0$ such that:

$$Q_k(x,\hat{y}_k) > Q_k(x,0) = 0 .$$

Noting that $P_u(x) = \sup_{0 \leq y \in \mathbb{R}^m} Q(x,y) > \sup_{0 \leq y_k \in \mathbb{R}} Q_k(x,y_k)$ \hfill (10)

and using the previous inequality, it follows that $P_u(x) > 0$ which proves the second part of (i).

(ii) Let $g_k(x) > 0$ for some $k \in I$. By (10) we have

$$P_u(x) > \sup_{y_k \geq 0} u^{-1} E u (y_k g_k(x,b)) .$$

Since u is increasing and $0 < g_k(x) \leq g_k(x,b)$, the above inequality implies that

$$P_u(x) > \sup_{y_k \geq 0} u^{-1} E u (y_k g_k(x)) = \sup_{y_k \geq 0} y_k g_k(x)$$

and hence the result (ii) follows. □

The first part of the theorem demonstrates that $P_u(x)$ is a penalty function for violation of the constraints in the mean. The second part shows that P_u has the desirable property of excluding solutions which are not feasible in (SP), for any realization of b, since for those $P_u(x) = \infty$.

We say that x^1 is less feasible than x^2, for the k-th constraint, if:

$$g_k(x^1,b) > g_k(x^2,b) \qquad \forall b \in B .$$

Note that for stochastic right hand side constraints $g_k(x) > b_k$, this simply means $g_k(x^1) < g_k(x^2)$.

4. PROPERTIES OF THE u-PENALTY

In this section we derive the basic properties of the u-penalty P_u, defined in the introduction (see eq. (4)), discuss the appropriateness of using it for programming problems with stochastic constraints, and derive an approximate simple expression for its computation.

<u>Theorem 2</u>   For any $u \in U_1$, the u-penalty function

$$P_u(x) = \sup_{y \geq 0} u^{-1} E\, u\, (y^T g(x,b))$$

satisfies:

(i) $P_u(x) = \begin{cases} 0 & \text{if } Eg(x,b) \leq 0 \\ \\ \text{positive} & \text{if } Eg(x,b) \nleq 0 \end{cases}$

(ii) $P_u(x) = \infty$ if for some k $g_k(x) = \inf_{b \in B} g_k(x,b) > 0$

<u>Proof:</u>   (i) Let $Q(x,y) = u^{-1} E\, u\, (y^T g(x,b))$ then

$$P_u(x) = \sup_{y \geq 0} Q(x,y) \geq Q(x,0) = 0 \qquad\qquad (9)$$

Since $u \in U_1$, by Jensen inequality (see Proposition 1 (c)):

$$Q(x,y) \leq y^T Eg(x,b)$$

with equality for $y = 0$, and so

$$P_u(x) = \sup_{y \geq 0} Q(x,y) \leq \sup_{y \geq 0} y^T Eg(x,b)$$

Now, if $Eg(x,b) \leq 0$, the last inequality shows that $P_u(x) \leq 0$ which together with (9) proves the first part of (i).

Assume now that for some $k \in I$, $Eg_k(x,b) > 0$.

Let $Q_k(x,y_k) := Q(x,0,\ldots y_k,\ldots,0) = u^{-1} E\, u\, (y_k g_k(x,b))$.

$$\frac{\phi E \, u \, (y^T Z) - \phi E \, u \, (x^T Z)}{\phi' E \, u \, (x^T Z)} \leq E \left(\frac{(y-x)^T Z}{\phi' (u(x^T Z))} \right) ; \quad \forall x, y \in \mathbf{R}^n \text{ and } Z. \quad (8)$$

The latter is equivalent to Jensen inequality for $h(t_1, t_2)$ with $t_1 = u(y^T Z)$ and $t_2 = u(x^T Z)$, $x, y \in R^n$, which holds for all Z if and only if h is convex. Thus, invoking lemma 1, the proof is completed.

□

Remark 1 - The two parameters class of utility function with hyperbolic absolute risk aversion (HARA) defined by $r(t) = \frac{1}{at+b}$, which is widley used in economics (see e.g., Hammond [1974], Wilson [1968]) satisfies trivially the condition that $\frac{1}{r(t)}$ is concave.

The (HARA) family consists of the following utilities (defined for $t > -\frac{b}{a}$):

$$u(t) = \begin{cases} -e^{-t/b} & \text{if } a = 0, \ b \neq 0 \\ \log(b+t) & \text{if } a = 1 \\ (at+b)^{(a-1)/a} & \text{if } a \neq 0, \ a \neq 1 \end{cases}$$

The first one, corresponding to constant risk aversion $r(t) = \frac{1}{b}$, and called accordingly (CRA)-utility function, is of particular interest and will be studied in Section 5.

where $\eta(y)$ is the <u>unique</u> solution of the equation:

$$E u'(y^T Z - \eta(y)) = 1 , \qquad \forall y \in \mathbb{R}^m . \qquad (24)$$

Moreover

(i) $\eta(y)$ (and hence $w(y)$) is continuously differentiable;

(ii) $\eta(0) = w(0) = 0; \quad \nabla\eta(0) = \nabla w(0) = E(Z).$

<u>Proof:</u> Let us define the function $\psi : \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}$ by:

$$\psi(y,\eta) = E u'(y^T Z - \eta) .$$

Since $u \in U_N$, $\psi(y,\eta)$ is continuously differentiable on $\mathbb{R}^{m+1}$.
Now equation (24) i.e. $\psi(y,\eta) = 1$ has a solution $\eta = \eta(y)$ for arbitrary
y. This follows since $\psi(y,\eta)$ is continuous in η, $\psi(y,\eta) < 1$ for η suf-
ficiently small and $\psi(y,\eta) > 1$ for η sufficiently large. Moreover since
$\frac{\partial}{\partial\eta} \psi(y,\eta) = -E u''(y^T Z - \eta) > 0$, by the implicit function theorem there exists
a unique differentiable solution $\eta = \eta(y)$ to the equation $\psi(y,\eta) = 1$. By the
definition of $w(y)$, in (22), as an unconstrained concave optimization
problem, $w(y)$ is obtained by equating the derivative of the supremum
to zero, thus (23) is proved.

Now (24) holds for any $y \in \mathbb{R}^m$, thus $\psi(0,\eta(0)) = u'(-\eta(0)) = 1$,
then, u' being strictly decreasing with $u'(0) = 1$, we get $\eta(0) = 0$

and hence from (23) $w(0) = 0$. Differentiating (24) with respect to y, gives for any $y \in \mathbb{R}^m$:

$$E((Z - \nabla \eta(y))u''(y^T Z - \eta(y))) = 0 .$$

Then, for $y = 0$, using the facts $\eta(0) = 0$ and $u''(0) < 0$, we have $\nabla \eta(0) = E(Z)$. Finally, differentiating (23) with respect to y we have using (24):

$$\nabla w(y) = E(Zu'(y^T Z - \eta(y))) \tag{25}$$

and thus $\nabla w(0) = E(Z)$.

$\square$

7. A PENALTY FUNCTION INDUCED BY THE NEW CERTAINTY EQUIVALENT

In analogy to the way the u-penalty was constructed in terms of the classical certainty equivalent (see eq. (4)), the new certainty equivalent induces the following penalty function:

$$\Pi_u(x) = \sup_{y \geqslant 0} S(y^T g(x,b)) = \sup_{y \geqslant 0} \sup_{\eta \in \mathbb{R}} \{\eta + E u (y^T g(x,b) - \eta)\} .$$

This function possesses the same properties as P_u, indeed from Theorem 4 (a), we see immediately that if b is deterministic then:

$$\Pi_u(x) = \begin{cases} 0 & g(x,b) \leqslant 0 \\ \\ +\infty & \text{otherwise} \end{cases}$$

thus it is an ideal penalty for the original deterministic program (P): $\inf\{g_0(x): g(x,b) \leqslant 0\}$. In the <u>stochastic</u> case it can be shown in analogy with Theorem 2 that :

Theorem 5 - The function Π_u satisfies :

$$\Pi_u(x) = \begin{cases} 0 & \text{if } Eg(x,b) \leqslant 0 \\ \text{positive and finite,} & \text{if } Eg(x,b) \nleqslant 0 \\ \infty & \text{if for some } k: g_k(x) > 0 \end{cases}$$

Proof: The proof follows step by step the proof of Theorem 2 since the latter uses properties of $C(x)$ which are shared by $S(x)$, see Theorem 4.

$\square$

Theorem 5 shows that Π_u is a penalty function for violation of constraints in the mean, and it automatically excludes solutions to the stochastic program (SP) which are not feasible for all realizations of b. Therefore a suitable deterministic surrogate problem for (SP) is:

$$\text{(NCE-P)} \quad \inf_{x \in \mathbb{R}^n} \{g_0(x) + \Pi_u(x)\} .$$

From the additivity of the new certainty equivalent (Theorem 4 (d)):

$$\inf_x \{g_0(x) + \Pi_u(x)\} = \inf_x \sup_{y \geqslant 0} S(g_0(x) + y^T g(x,b))$$

hence (NCE-P) can be written as a minimax problem :

$$\text{(NCE-P)} \quad \inf_x \sup_{y \geqslant 0} S(L_b(x,y))$$

where L_b is the Lagrangian of problem (SP).

We now derive an upper bound for the penalty Π_u in terms of the probability of satisfying the constraints: $Pr(g(x,b) \leqslant 0)$.

Theorem 6 - Let $u \in U_N^A$. For every x such that $p(x) = Pr\{g(x,b) \leqslant 0\} > 0$, one has:

$$\Pi_u(x) \leqslant \psi(p(x)) \qquad (27)$$

where $\psi(p)$ is given in Lemma 2.

Proof: Fix x such that $p(x) > 0$. Recall that:

$$\Pi_u(x) = \sup_{y \geqslant 0} S(y^T g(x,b)) .$$

Applying Lemma 2 with $X = y^T g(x,b)$ for some non zero $y \geqslant 0$, and $L = 0$, we get

$$S(y^T g(x,b)) \leqslant \psi(P_y) \qquad (28)$$

where here $P_y = P_y(x) = Pr(y^T g(x,b) \leqslant 0)$. Hence from (28)

$$\Pi_u(x) \leqslant \sup_{y \geqslant 0} \psi(P_y) \qquad (29)$$

Moreover, since ψ is decreasing (see Lemma 2):

$$\sup_{y \geqslant 0} \psi(P_y) \leqslant \psi(\inf_{y \geqslant 0} Pr(y^T g(x,b) \leqslant 0)). \qquad (30)$$

But for all $0 \neq y \geqslant 0$: $Pr(y^T g(x,b) \leqslant 0) \geqslant Pr(g(x,b) \leqslant 0)$, and hence, using again the fact that ψ is decreasing,

$$\psi(\inf_{y \geqslant 0} Pr(y^T g(x,b) \leqslant 0)) \leqslant \psi(Pr(g(x,b) \leqslant 0)$$

which combined with (29) and (30) proves the inequality (27).

□

Remark 2: Using the properties of the function $\psi(\cdot)$ from Lemma 2, it is interesting to note that the larger is the probability of satisfying the constraints, the smaller is the "upper-bound penalty" $\psi(p(x))$.

We derive now, quadratic approximation for the penalty function $\Pi_u(x)$. We use the same notations as given in Section 4. Let us denote by $w_x(y)$ the NCE-functional corresponding to the random vector $Z = g(x,b)$. From the Proposition 6, for $u \in U_N$ the NCE

functional is given by:

$$w_x(y) = n_x(y) + E u (y^T g(x,b) - n_x(y)) \tag{23}$$

with $n_x(y)$ uniquely determined by $E u' (y^T g(x,b) - n_x(y)) = 1$.

A direct differentiation of (23) with respect to y gives:

$$\nabla w_x(0) = m(x)$$

$$\nabla^2 w_x(0) = u''(0)\Sigma(x) = - r_o \Sigma(x)$$

(since here $u'(0) = 1$, $r_o = - \dfrac{u''(0)}{u'(0)} = - u''(0)$).

Moreover, we know that $w_x(0) = 0$, hence a second order Taylor expansion of $w_x(y)$ is:

$$\hat{w}_x(y) = y^T m(x) - \frac{1}{2} r_o y^T \Sigma(x) y$$

and therefore an approximation of $\Pi_u(x)$ is given as the optimal value of the concave quadratic program:

$$\hat{\Pi}_u(x) = \sup_{y \geqslant 0} \{y^T m(x) - \frac{1}{2} r_o y^T \Sigma(x) y\} .$$

For a quadratic utility, the approximation is exact; this can be verified by a direct calculation. The approximation $\hat{\Pi}_u(x)$, coincides with the approximate $\hat{P}_u(x)$ of the u-penalty given in Section 4. In particular, similar results concerning uncorrelated constraints (see Proposition 3) can be recovered here, i.e.:

$$\hat{\Pi}_u(x) = \frac{1}{2r_o} \sum_{i=1}^{m} \frac{1}{\sigma_i^2(x)} [\max(0,m_i(x))]^2 .$$

8. DUALITY RESULTS FOR STOCHASTIC RHS PROGRAMS

In this section we treat the special case of the general stochastic programming problem (SP):

(SP-RHS) $\inf\{g_0(x): g_i(x) \geqslant b_i, \quad i \in I = 1,\ldots,m\}$

which is simply obtained from (SP) with $g(x,b) = b - g(x)$. For each $i \in I$, $\underline{b}_i$ will denote the infinum support of b_i. Using the new certainty equivalent S, the penalty function Π_u is given here by:

$$\Pi_u(x) = \sup_{y \geqslant 0} S(y^T(b-g(x))) \quad .$$

However, by the additivity of S, since $y^T g(x)$ is not random, we obtain the following representation for Π_u :

$$\Pi_u(x) = \sup_{y \geqslant 0} \{w(y) - y^T g(x)\}$$

where as in Section 6 :

$$w(y) = S(y^T b) = \sup_{\eta \in \mathbb{R}} \{\eta + \mathbb{E}\, u\,(y^T b - \eta)\} \quad .$$

The corresponding deterministic primal (NCE-P): $\inf \{g_0(x) + \Pi_u(x)\}$ is then:

(NCE-P) $\inf\limits_{x} \{g_0(x) + \sup\limits_{y \geqslant 0} \{w(y) - y^T g(x)\}\} \quad .$

Assume from now on that $g_0(x)$ is convex, and that $\{g_i(x)\}_{i \in I}$ are concave functions, so (SP-RHS) is a convex program. This implies that $\Pi_u(x)$ is convex and so (NCE-P) is a convex program. In terms of the saddle function:

$$K(x,y) = g_0(x) + w(y) - y^T g(x) \tag{31}$$

it can be written as:

(NCE-P) $\inf_{x} \sup_{y \geqslant 0} K(x,y)$.

We define the Dual problem corresponding to (NCE-P) by

(NCE-D) $\sup_{y \geqslant 0} \inf_{x} K(x,y)$.

The main result concerning the dual pair (NCE-P) and (NCE-D) is the validity of a <u>strong duality</u> result:

$$\inf (NCE\text{-}P) = \max (NCE\text{-}D) \quad .$$

This will be proved in Theorem 8. Before we need an additional limit property of the (one dimensional) NCE function $w(y)$, which is interesting by itself. A similar property for the classical certainty equivalent functional $v(y)$, but <u>only</u> for the exponential utility is given in Bamberg and Spremann [1981].

<u>Theorem 7</u> - Let X be a random variable, with $x_{min} > -\infty$ denoting the infimum of the support of X, and let $u \in U_N^A$. Then

$$\lim_{y \to +\infty} \frac{w(y)}{y} = x_{min} \qquad (32)$$

<u>Proof</u>: We have to prove that

$$\lim_{y \to +\infty} \frac{w(y)}{y} > x_{min} \qquad (33)$$

$$\lim_{y \to +\infty} \frac{w(y)}{y} \leqslant x_{min} + \epsilon \qquad \forall \epsilon > 0 \qquad (34)$$

For any $y > 0$, $yX-\eta \geqslant y\,x_{min}- \eta$ and since u is increasing, from the definition of $w(y)$ we get:

$$w(y) \geqslant \sup_{\eta \in \mathbb{R}} \{\eta + u(y\,x_{min}-\eta)\} = S(y\,x_{min}) \quad .$$

By Theorem 4 (a), $S(y\,x_{min}) = y\,x_{min}$, thus the last inequality proves (33).

Let $\epsilon > 0$ be fixed, but arbitrary. Applying Lemma 2 to the random variable yX with $L = y(x_{min} + \epsilon)$, we have :

$$w(y) = S(yX) \leqslant y(x_{min} + \epsilon) + \psi(p)$$

where $p = Pr(X \leqslant x_{min} + \epsilon) > 0$. From the latter inequality, (34) follows immediately.

$\square$

The duality theorem now follows.

Theorem 8 - Let (SP-RHS) be a convex stochastic program and consider the corresponding deterministic program (NCE-P) for $u \in U_N^A$.

If the following condition holds :

(S) $\exists \hat{x} \in \mathbb{R}^n$ such that $g_i(\hat{x}) > \underline{b}_i$ $\forall i \in I$

then

$$\inf (NCE\text{-}P) = \max (NCE\text{-}D) \tag{35}$$

Proof: Since $g_0(x)$ is convex and $\{g_i(x)\}_{i \in I}$ are concave, then $K(\cdot,y)$ given in (31) is convex for every $y \geqslant 0$. From Proposition 5 , we know that $w(y)$ is concave, hence $K(x,\cdot)$ is concave. By a result of Rockafeller [1964] a sufficient condition for the validity of (35), for a general convex-concave saddle function $K(x,y)$ is:

$$\exists\ 0 \neq y_0 \geqslant 0 \text{ such that } y_0^T \nabla_y K(x,y) \geqslant 0 \quad (x \in \mathbb{R}^n, y > 0) \quad .$$

Here by Proposition 6, (see eq. (25)):

$$\nabla_y K(x,y) = \nabla w(y) - g(x) = E(bu'(y^t b - \eta(y))) - g(x)$$

thus, we have to show :

$$\exists y_o \geq 0 \text{ such that } y_o^T \{E(bu'(y^T b - n(y))) - g(x)\} \geq 0$$
$$\forall x \in \mathbb{R}^m, \ y > 0 \ .$$

This is certainly satisfied if:

$$\exists \hat{x}, \ \exists \hat{y} > 0 \text{ such that } \quad \nabla w(\hat{y}) = E(bu'(\hat{y}^T b - n(\hat{y}))) < g(\hat{x}) \ . \qquad (36)$$

To show that condition (S) implies (36) it suffices to prove that:

$$\inf_{y \geq 0} \frac{\partial}{\partial y_i} w(y) \leq \underline{b}_i \qquad i \in I \qquad (37)$$

For all $i \in I$, let $w_i(y_i) = w(0,0,\ldots,y_i,\ldots 0)$ i.e.,

$$w_i(y_i) = \sup_{\eta \in \mathbb{R}} \{\eta + Eu(b_i y_i - \eta)\} \ .$$

Now,

$$\inf_{0 \leq y \in \mathbb{R}^m} \frac{\partial}{\partial y_i} w(y) \leq \inf_{0 \leq y_i \in \mathbb{R}} \frac{\partial}{\partial y_i} w_i(y_i) \qquad \forall i \in I \ .$$

So, in order to prove (37) it suffices to prove that

$$\inf_{0 \leq y_i \in \mathbb{R}} w_i'(y_i) \leq \underline{b}_i \ . \qquad (38)$$

But w_i' is a derivative of a strictly concave function and thus is strictly decreasing, hence $\inf_{y_i \geq 0} w_i'(y_i) = \lim_{y_i \to \infty} w_i'(y_i)$.

Moreover, $w_i(y_i)$ is concave, thus by the gradient inequality

$$0 = w_i(0) \leq w_i(y_i) - y_i w_i'(y_i)$$

and then by Theorem 7

$$\lim_{y_i \to +\infty} w_i'(y_i) \leqslant \lim_{y_i \to +\infty} \frac{w_i(y_i)}{y_i} = b_i$$

which is exactly (38), and the proof is completed.

□

Remark 3 -   The regularity condition (S) needed to establish the strong duality result for the pair of problems (NCE-P) (NCE-D) is extremely mild; for continuous random variables, if (S) does not hold, then the feasible set $\{x: g(x) \geqslant b\}$ is empty in probability 1.

□

Our last result gives a concrete interpretation of the dual problem (NCE-D). Recall that in the deterministic case the Lagrangian dual of (P) is the concave program: $\sup_{y \geqslant 0} [\inf_x L_b(x,y)]$. We show here, in the stochastic case, that the dual program (NCE-D) consists of maximizing the new certainty equivalent of the Lagrangian dual function.

Theorem 9 - Let (SP-RHS) be a convex stochastic program, and $u \in U_N$. Then the induced deterministic dual (NCE-D) is the concave program:

$$\text{(NCE-D)} \quad \sup_{y \geqslant 0} S(\inf_x L_b(x,y)) \tag{39}$$

Proof: The dual problem (NCE-D) is given by

$$\text{(NCE-D)} \quad \sup_{y \geqslant 0} \inf_x K(x,y)$$

where $K(x,y)$ is defined in (31). Since $g_o(x) - y^T g(x)$ is not random, by the additivity property of $S(\cdot)$, (Theorem 4 (d)) $K(x,y)$ can be rewritten as :

$$K(x,y) = S(L_b(x,y)) \tag{40}$$

where $L_b(x,y) = g_o(x) + y^T(b - g(x))$ is the Lagrangian for (SP-RHS).

Thus, in order to prove (39), we need to prove that

$$\inf_x S(L_b(x,b)) = S(\inf_x L_b(x,y))$$

Indeed, we have:

$$
\begin{aligned}
S(\inf_x L_b(x,y)) &= S(y^Tb + \inf_x(g_o(x) - y^Tg(x))) \\
&= S(y^Tb) + \inf_x(g_o(x) - y^Tg(x)) \quad \text{[by Theorem 4 (d)]} \\
&= \inf_x\{S(y^Tb) + g_o(x) - y^Tg(x)\} \\
&= \inf_x S(y^Tb + g_o(x) - y^Tg(x)) \quad \text{[by Theorem 4 (d)]} \\
&= \inf_x S(L_b(x,y))
\end{aligned}
$$

$\square$

REFERENCES

1. Arrow, K.J., *Essays on the Theory of Risk-Bearing*, Markham Publishing, Chicago, 1971.

2. Bamberg, G. and Spremann, K., "Implications of Constant Risk Aversion", *Zeit. fur Operations Research*, Vol. 25 (1981), pp. 205-224.

3. Ben-Tal, A., "The Entropic Penalty, Approach to Stochastic Programming", to appear in *Math. of Oper. Res.*

4. Ben-Tal, A. and Teboulle, M., "The Duality Between Expected Utility and Penalty in Stochastic Linear Programming", Proceedings of the IFIP Conf. on "Stochastic Programming", F. Archetti (ed.), Lecture Note in Control and Information Sciences, Springer Verlag, NY., 1984, to appear.

5. Bourbaki, N., *Elements de Mathematique, Fonctions d'une Variable Reelle*, Vol. IX, Livre IV, Hermann, Paris, 1958.

6. Dantzig, G.B., "Linear Programming under Uncertainty", *Management Sc.*, Vol. 1 (1955), pp. 197-206.

7. Dantzig, G.B., and Madansky, A., "On the Solution of Two-Stage Linear Programs under Uncertainty", in *Proc. Fourth Berkeley Sympos. on Mathematical Statistics and Probability*, Neyman, I.J. (ed.), 1961, pp. 165-176.

8. Dempster, M.A.H., (ed.) *Stochastic Programming*, Academic Press, New-York, 1980.

9. Fiacco, A.V., and McCormick, G.P., *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, John Wiley, New-York, 1968.

10. Hammond, J.S., "Simplifying the Choice between Uncertain Prospects where Preference is Non-Linear", *Management Sci.*, Vol. 20 (1974), pp. 1047-1072.

11. Kall, P., "Stochastic Programming", Invited Review, *European Journal of Oper. Res.*, Vol. 10 (1982), pp. 125-123.

12. Luenberger, D.G., *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, Reading, Mass., 1973.

13. Mangasarian, O.L. and Rosen, J.B., "Inequalities for Stochastic Nonlinear Programming Problems", *Operations Res.*, Vol. 12 (1964), pp. 143-154.

14. Mood, A.M., Graybill, F.A., and Boes, D.C., *Introduction to the Theory of Statistics*, McGraw Hill Series in Probability and Statistics, 1974.

REFERENCES, con't.

15.   Pratt, J.W., "Risk Aversion in the Small and in the Large,"
        Econometrica, Vol. 32, (1964), pp. 122-136.

16.   Rockafellar, R.T., "Minimax Theorems and Conjugate Saddle Functions,"
        Math. Scand., Vol. 14, (1964), pp. 151-173.

17.   Rockafellar, R.T., Convex Analysis, Princeton University Press,
        Princeton, N.J., 1970.

18.   Rockafellar, R.T., Conjugate Duality and Optimization, Regional
        Conference Series in Applied Mathematics, No. 16, SIAM, 1974.

19.   von Neuman, J. and Morgenstern, O., Theory of Games and Economic
        Behavior, 2nd ed., Princeton University Press, Princeton, N.J., 1967.

20.   Walkup, D. and Wets, R., "Stochastic Programs with Recourse,"
        SIAM J. Appl. Math., Vol. 15, (1967), pp. 1299-1314.

21.   Wets, R., "Stochastic Programming: Solution Techniques and
        Approximation Schemes", Mathematical Programming The State of the
        Art, A. Bachem, M. Grotschel, B. Korte (eds.), Springer-Verlag,
        Berlin, 1983.

22.   Wilson, R., "The Theory of Syndicates", Econometrica, Vol. 36,
        (1968), pp. 119-132.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER CCS 506	**2. GOVT ACCESSION NO.** *AD.A154034*	**3. RECIPIENT'S CATALOG NUMBER**
4. TITLE *(and Subtitle)* Expected Utility, Penalty Functions and Duality in Stochastic Nonlinear Programming		**5. TYPE OF REPORT & PERIOD COVERED**
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(*s*) A. Ben-Tal M. Teboulle		**8. CONTRACT OR GRANT NUMBER(*s*)** N00014-82-K-0295
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Cybernetic Studies The University of Texas at Austin Austin, Texas 78712		**10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS**
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research (Code 434)		**12. REPORT DATE** February 1985
		13. NUMBER OF PAGES 38
14. MONITORING AGENCY NAME & ADDRESS(*if different from Controlling Office*)		**15. SECURITY CLASS. *(of this report)*** Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Stochastic Programming, Duality in Nonlinear Programming, Minmax Theorems
Convex Functions, Expected Utility, Risk Aversion, Certainty Equivalent,
Penalty Functions

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

We consider nonlinear programming problems with stochastic constraints.
The Lagrangian corresponding to such problems has a stochastic part, which
in this work is replaced by its certainty equivalent (in the sense of
expected utility theory). It is shown that the deterministic surrogate
problem thus obtained, contains a penalty function which penalized
violation of the constraints in the mean. The dual problem is studied
(for problems with stochastic righthand sides in the constraints) and a

Abstract continued.

comprehensive duality theory is developed by introducing a new certainty
equivalent concept, which possesses, for arbitrary utility functions, some
of the properties that the classical certainty equivalent retains only
for the exponential utility.

END

FILMED

6-85

DTIC